AD_____

Award Number: DAMD17-00-1-0353

TITLE: Microarray Technology to Study the Role of Genetic Polymorphisms in Breast Cancer Risk

PRINCIPAL INVESTIGATOR: Hilmi Ozcelik, Ph.D.
Julia A. Knight

CONTRACTING ORGANIZATION: Mount Sinai Hospital
Toronto Ontario M5G 1X5 Canada

REPORT DATE: July 2002

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 074-0188*



| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | July 2002 | Annual (1 Jun 01 - 1 Jun 02) |

**4. TITLE AND SUBTITLE**
Microarray Technology to Study the Role of Genetic Polymorphisms in Breast Cancer Risk

**5. FUNDING NUMBERS**
DAMD17-00-1-0353

**6. AUTHOR(S)**
Hilmi Ozcelik, Ph.D.
Julia A. Knight

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Mount Sinai Hospital
Toronto Ontario M5G 1X5 Canada

E-Mail: Ozcelik@mshri.on.ca

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**
Report contains color

**12a. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for Public Release; Distribution Unlimited

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 Words)*

Several studies have suggested an association between low penetrant alleles and breast cancer risk. Although the contribution of low penetrant alleles to the individual risk is relatively small, they can contribute to a large proportion of breast cancer case in the population. In this study we took the candidate gene approach to study the association of 32 different genetic polymorphisms with breast cancer risk in a population-based sample using a high-throughput SNParray technology. To date, we have successfully established the SNParrag method. We have evaluated the specificity of individual probes to detect genotypes correctly, and developed a multiplex hybridization procedure. We have also developed a complementary high-throughput genotyping method (TaqMan) to study SNPs which did not agree with SNParray method. We have developed a software to organize and evaluate the raw data obtained from SNParray scans. We have prepared and plated all the breast cancer and population control DNA samples to be studied. We have validated the SNParray method by comparing it to two other genotyping methods. Using the established technologies we are currently in progress of completing the genotyping of the proposed breast cancer cases and population controls.

**14. SUBJECT TERMS**
cancer genetics, predisposition, SNP, polymorphism, microarray, SNParray, low penetrating alleles, breast cancer

**15. NUMBER OF PAGES**
20

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | Unlimited |

## Table of Contents

## INTRODUCTION

It has long been hypothesized that genetic variation is responsible for observed differences in cancer risk and susceptibility amongst the human population. Mutant alleles of dominant highly penetrant breast cancer genes, including BRCA1 and BRCA2 (1-3), do not occur frequently, and hence account for only a small proportion of breast cancer cases. On the other hand, several studies have suggested an association between low penetrant alleles and breast cancer risk. Although the contribution of low penetrant alleles to the individual breast cancer risk is relatively small, they can contribute to a large proportion of breast cancer cases in the population because the risk-conferring alleles of these genes are common.

Identification and cloning of low penetrant alleles that increase the risk of breast cancer is challenging because the association methods for such studies require large populations to achieve meaningful statistical analysis and very dense genetic maps to facilitate genome-wide genotyping (4,5). Although microarray technology has been developed to the point where it could be applied for parallel analysis of genome-wide genotyping (6), the dense genetic maps required for large population based association studies are currently being constructed for future genome-wide applications and will not be available for several years. At present, the candidate gene approach remains the most logical and practical strategy to identify these risk enhancing, low penetrant variants or single nucleotide polymorphisms (SNPs). Until now, a major obstacle with investigating the risk associated with multiple candidate genes has been a lack of technology for large-scale genotyping of large populations. Consequently, many studies have focused efforts on only 1 or 2 genetic polymorphisms, and even in these cases the analysis was only limited to relatively small sample sizes. In the context of the ideas program, we propose to exploit the high throughput power of SNParrays to simultaneously genotype 31 different genetic polymorphisms derived from 26 genes in a well defined, representative population-based sample containing a large number of subjects. We have selected genetic polymorphisms in genes involved in different aspects of carcinogenesis (7-41). For example, cell cycle regulatory genes such as CDK-inhibitors, and cyclins; carcinogen metabolizing enzymes such as CYPs, GSTs and NATs; immune system genes such as interluekins and TNF; and genes involved in other pathways involved in cancer (e.g. p53, PTEN, XPD-DNA repair gene). We have access to the Ontario Familial Breast Cancer Registry (OFBCR), which is the largest population based breast cancer registry in Canada. We also have support from the established microarray facilities of the Ontario Cancer Institute and Samuel Lunenfeld Research Institute in Toronto.

The main objective of the proposed work is to identify low penetrant, yet commonly occurring, genetic polymorphisms which contribute to the risk of developing breast cancer. Furthermore, this approach has the potential to identify novel genetic factors associated with breast cancer risk, which may result in the development of innovative therapies, and a fuller understanding of genetic variation in response to therapy. The establishment of the proposed approach will prepare us for large-scale genotyping involving hundreds or even thousands of candidate genes in large define populations. This will lead to a more complex analysis of gene-gene and gene-environment interactions than is currently possible. Advances in disease etiology will significantly expand our abilities to design strategies for the prevention of breast cancer development and progression.

## STATEMENT OF WORK

Task 1: Characterization of polymorphic alleles by SSCP, Months 1-8

a. Design of SSCP primers for 32 sites
b. Screen by SSCP analysis for all possible alleles at each locus
c. Sequence the SSCP patterns (appr 3 per loci) and identify the all possible genotypes

Task 2: Designing of oligonucleotides and sample microarrays, Months 4-8

a. Design different sets of oligonucleotides ( perfect matches and mismatches)
b. Customize sample chips for quality control of hybridizations

Task 3: Optimization of the hybridizations using PCR probes, Months 8-16

a. Prepare PCR probes using control specimens
b. Optimize the hybridization conditions
c. Evaluate the accuracy of detection for every polymorphic site using a probes with different allelic combinations for each polymorphism
d. Redesign oligonucleotides and chips in order to increase the quality and accuracy of detection

Task 4: Genotyping of 900 specimens for 32 polymorphisms, Months 16-32

a. Production of microarray chips
b. Preperation of flourescent labelled PCR probes for each patient
c. Hybridization of chips at optimized conditions
d. Reading and analysis of the chip signals
e. Quality control experiments at different intervals using the control specimens to ensure the reproducibility of results

Task 5: Data and statistical analysis, Months 32-36

a. Repeat and conformation experiments
b. Complete the reading of every slide and prepare the data for statistical analysis
c. Univariate analysis of the data
d. Exploratory multivariate analysis of the data

## BODY

**Design and Production of SNParrays (Task 3 and Task 4)**

Support oligonucleotides are designed to be printed on SNParrays and bind allele specific probes. Each support-oligonucleotide contains an anti-TAG sequence which is complementary to the TAG-sequences on each of the allele-specific oligonucleotide sequences. Each anti-TAG-sequence is also attached to a 15-mer poly (T)-tail. The anti-TAG sequences on the support-oligonucleotides are designed to hybridize the fluorescently labeled allele-specific oligonucleotides, thus attaching them to the designated spots on the glass slides. The poly (T)-tails, on the other hand, are designed to increase the efficiency of support-oligonucleotides to bind to the glass surface during printing.

Before printing, support-oligonucleotides are diluted to 100μM concentration in 3xSSC buffer and placed in 384 well V-bottom distribution plates. The arrayer is programmed to print these support-oligonucleotides according to the design given below (Figure 1). Each support-oligonucleotide is printed in duplicate for validation purposes.

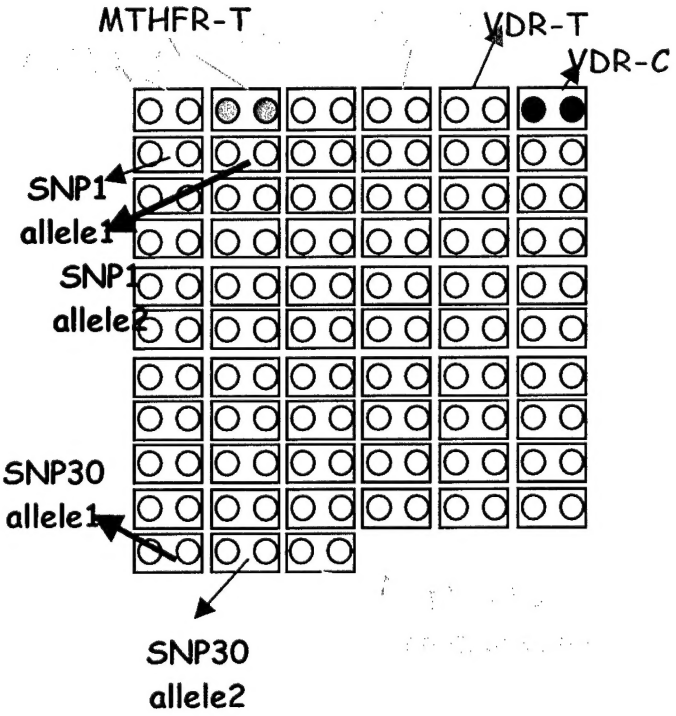



**Figure 1**: Design of the SNP arrays. Each support oligonucleotide is printed in duplicate.

The SNParrays are printed on slides with poly-L-lysine surface chemistry by the microarray facility of Samuel Lunenfeld Research Institute (SLRI) of Mount Sinai Hospital (MSH) in Toronto. The mechanism of poly-L-Lysine slides to bind oligonucleotides is very similar to the super-amine slides. The slides are covered with a positively charged chemical, Lysine in this case, for the negatively charged oligonucleotides to bind where the poly (T)-tail enhances the strength of binding. A Virtek•ChipWriter™ Professional (Pro) arrayer is used for printing spots with a diameter of 120μM. The distance between the centers of two adjacent spots is 200μM. Each support oligonucleotide is placed in two different array locations on the distribution plates, and two pins are used to remove oligonucleotides from these two locations and print duplicate arrays on the same slide simultaneously (Figure 2).

The spots on the slides are rehydrated after printing process. A humidity chamber is filled with 100ml of 1xSSC solution and the slides are placed in the chamber, and kept there for 60 seconds (array side is facing to the solution). The arrays are then snap-dried for a few seconds on a $100^0$C hot plate (array side facing up). Rehydration process increases the exposure of oligonucleotides to probes during hybridization. After printing, oligonucleotides on the slide are fixed in a UV cross-linker at 600mJ. Unbound oligonucleotides and excess salt were washed off the slides for 15 minutes at room temperature in a solution of 170mM succinic anhydride (in a solution of 1,2 methyl pyrollidinone and 40mM boric acid). This step neutralizes the positively charged surface of the slide to limit the hybridization of the negatively charged probe to the support-oligonucleotides only, and not to the slide surface. The slides are then rinsed in distilled water for one minute, and then kept in 95% ethanol for an additional minute. They are then spin-dried at 600 rpm for 10 minutes.

The printing quality of SNParrays is controlled with a Cy3 labeled control oligonucleotide hybridization. Synthesized poly (A)-tail probes are attached to Cy3, and hybridized under the hybridization conditions described below. The quality of every new batch of printed slides was validated using this approach (Figure 2).

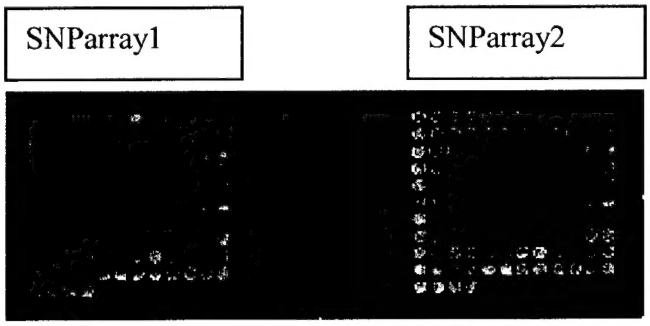



**Figure 2**: Quality control of printed oligonucleotides. The duplicate SNParrays are hybridized with cy3-labelled poly (A) probes. Homogenous signal is detected along the array. The last pair of control oligonucleotides (non-poly (T)) did not give any signal.

**Optimization of Probe Specificity on SNParrays (Task 3)**

PCR reactions for probe preparation purposes were optimized taking into consideration different variables that affected the yield and the specificity of the product.

Reactions are performed in a total volume of 10µl, in the presence of 10µM of each cold dNTP, a range of 2-4mM $MgCl_2$, 5pmol of each allele specific primer, 10pmol of the common reverse primer, 5µM of fluorescently (cy5) labeled dCTP, 0.25U of Platinum Taq polymerase, and 10ng genomic DNA. Different annealing temperatures are tested (ranging from 55-65$^0$C), depending on the melting temperatures of the PCR primers used. For the quality control purposes, previously known homozygote and heterozygote templates are used to prepare the probes for each SNP. Each probe is hybridized to duplicate arrays to ensure that the genotypes were detected correctly. This procedure is repeated for all the SNPs in the study.

Initial quality control experiments have revealed that only 10 of 31 SNPs could be detected specifically for the combination of three allelic statuses. An example of an unspecific SNP detection is given for IFA13 SNP in Figure 3. Experiments for the unspecific SNPs were repeated using a wider range of $MgCl_2$ and annealing temperature conditions. With this we were able to increase the specificity and recover five additional SNPs. New set of allele-specific primers were designed for the remaining unsuccessful SNPs, and the optimizations were carried out as described above. We were able to obtain specific detection from an additional of four SNPs. As a conclusion, we were able to obtain high specificity from 19 out of 31 SNPs in the study. See Table 1 for the working conditions of these SNPs. The unspecific hybridization of the probes was mainly resulted from incompatibility between the nature of the sequences around the corresponding SNPs and the experimental design for probe preparation for SNP arrays. We have developed an alternative complementary method to screen for these SNPs successfully (see TaqMan method below).

**Multiplex Hybridization and Signal Detection (Task 3 and 4)**

The PCR reactions (probes) are prepared for each SNP using 96-well microplates and kept frozen until all the SNPs for an individual DNA is completed. After that, 5µl of each PCR was pooled in another plate with an 8-channel micro-pipette, and mixed. From the pooled probe mixture, 15µl was removed and mixed with 5 µl of a hybridization mixture (1.33xSSC, 0.067% SDS, 0.033mg/ml of salmon sperm DNA). This mixture is placed over the SNParray which is then covered with a cover-slip. The slides are incubated for 3 hrs at 50$^0$C, after which they are washed for 15 minutes in a solution of 2xSSC and 0.1% SDS. SDS is rinsed off with two additional consecutive washings in 2xSSC, 15 minute each. Slides were dried in a centrifuge at 600 rpm for 10 minutes. Slides are kept in a dark and dry place to eliminate the fluorescence bleaching due to light exposure.

Detection of the intensity of fluorescent signal after hybridization is performed in a GenePix 4000B slide scanner (Axon). The scan results are then analyzed with a Genepix Pro 4.0 analysis software. The software uses a feature indicator which sits on the SNParray. Each feature in this block defines the borders of individual oligonucleotide spots. Anything outside the border (until the next spot) is considered as background, and anything inside the border is considered as gross signal intensity. By creating a special type of tab delimited text file (.gal), oligonucleotide names can be defined to the individual features. This file is created with the names of oligonucleotides, and their specific locations are defined in the SNParray. A results table is created by the software, including the medians and means of the gross signal intensity, background signal intensity and the net signal intensity. This table is saved as a tab delimited text file (.gpr extension), which can be opened and examined in a Microsoft Excel worksheet format. The means of net signal intensity (after subtracting the background from the gross signal) is used for determination of the genotypes. An example of raw (Table 2) and processed data (after subtraction of background values) (Figure 4) are given.

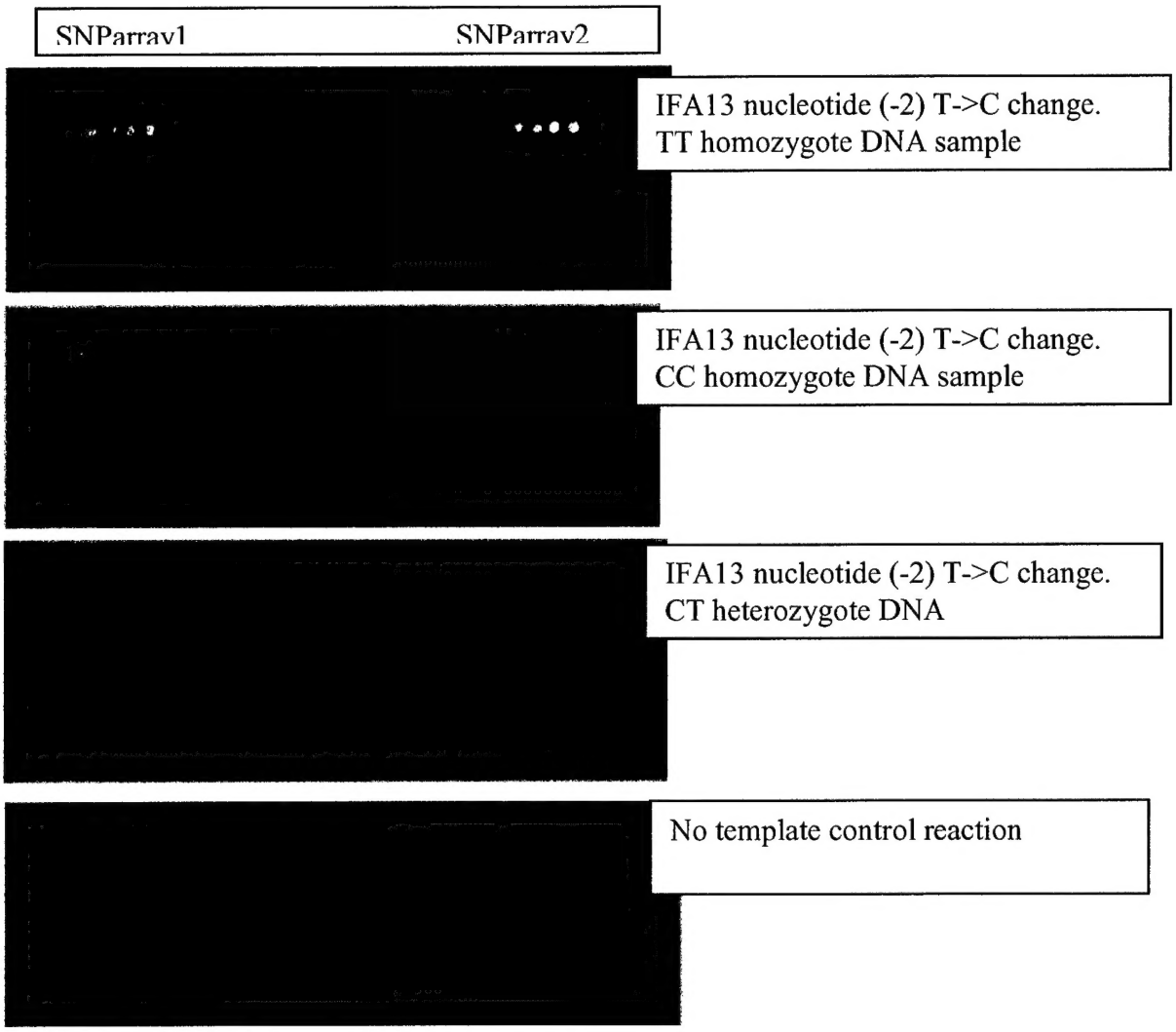


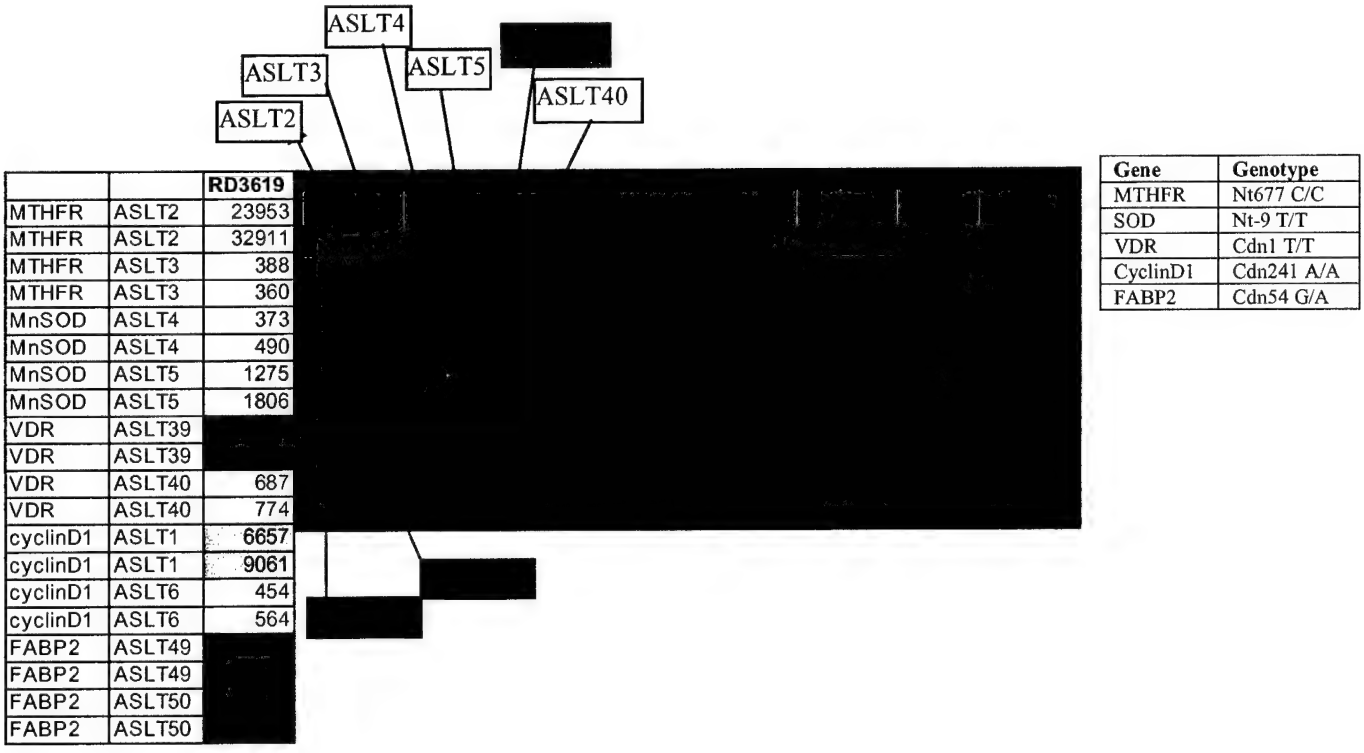

**Figure 3**: Measuring the specificity and the quality of each SNP specific probe to detect allelic status. IFA13 SNP was studied using probes prepared from homozygote (CC and TT) and heterozygote (CT) templates. As seen in the figure the genotype detection was not specific to the allelic status.

| gene | primer | Mg | dNTP | primer (pmole) | T(oC) |
|---|---|---|---|---|---|
| MTHFR | ASLT2 | 4 | 10 | 5 | 56 |
| MTHFR | ASLT3 | 4 | 10 | 5 | 56 |
| MnSOD | ASLT4 | 2 | 10 | 5 | 56 |
| MnSOD | ASLT5 | 2 | 10 | 5 | 56 |
| VDR | ASLT39 | 4 | 10 | 5 | 60 |
| VDR | ASLT40 | 4 | 10 | 5 | 60 |
| cyclinD1 | ASLT1 | 4 | 10 | 5 | 68 |
| cyclinD1 | ASLT6 | 4 | 10 | 5 | 68 |
| p27/kip1 | ASLT11 | 2 | 10 | 1.5 | 63 |
| p27/kip1 | ASLT12 | 2 | 10 | 1.5 | 63 |
| p53 | newASLT17 | 6 | 20 | 5 | 60 |
| p53 | newASLT18 | 6 | 20 | 5 | 60 |
| GADD45 | ASLT19 | 4 | 20 | 5 | 60 |
| GADD45 | ASLT20 | 4 | 20 | 5 | 60 |
| ERcdn10 | ASLT21 | 4 | 10 | 5 | 60 |
| ERcdn10 | ASLT22 | 4 | 10 | 5 | 60 |
| ERcdn325 | ASLT25 | 4 | 20 | 5 | 56 |
| ERcdn325 | ASLT26 | 4 | 20 | 5 | 56 |
| MMP-1 | ASLT27 | 4 | 10 | 5 | 56 |
| MMP-1 | ASLT28 | 4 | 10 | 5 | 56 |
| BARD1 | ASLT31 | 4 | 20 | 5 | 56 |
| BARD1 | ASLT32 | 4 | 20 | 5 | 56 |
| XPD | ASLT33 | 4 | 10 | 5 | 60 |
| XPD | ASLT34 | 4 | 10 | 5 | 60 |
| GSTM3 | ASLT37 | 4 | 10 | 5 | 60 |
| GSTM3 | ASLT38 | 4 | 10 | 5 | 60 |
| CYP17 | newASLT45 | 4 | 10 | 5 | 65 |
| CYP17 | newASLT46 | 4 | 10 | 5 | 65 |
| COMT | ASLT47 | 4 | 20 | 5 | 56 |
| COMT | ASLT48 | 4 | 20 | 5 | 56 |
| FABP2 | ASLT49 | 4 | 10 | 5 | 56 |
| FABP2 | ASLT50 | 4 | 10 | 5 | 56 |
| TNF-a | ASLT51 | 4 | 10 | 5 | 68 |
| TNF-a | ASLT52 | 4 | 10 | 5 | 68 |
| G-CSF | ASLT57 | 4 | 10 | 5 | 68 |
| G-CSF | ASLT58 | 4 | 10 | 5 | 68 |
| IL-13 | ASLT59 | 4 | 10 | 5 | 56 |
| IL-13 | ASLT60 | 4 | 10 | 5 | 56 |

**Table 1**: PCR (probe preparation) conditions for 19 SNPs detection of which were specific on SNParrays.

| Block | Column | Row | Name | ID | X | Y | Dia. | F635 Median | F635 Mean | F635 SD | B635 Median | B635 Mean | B635 SD | Sum of Medians | Sum of Means | F635 Median - B635 | F635 Mean - 3635 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | MTHFR | ASLT2 | 8850 | 36510 | 120 | 47761 | 43541 | 14102 | 745 | 815 | 450 | 47016 | 42796 | 47016 | 42796 |
| 1 | 2 | 1 | MTHFR | ASLT2 | 9060 | 36510 | 120 | 50374 | 45463 | 17465 | 754 | 791 | 412 | 49620 | 44709 | 49620 | 44709 |
| 1 | 3 | 1 | MTHFR | ASLT3 | 9250 | 36510 | 120 | 1181 | 1173 | 533 | 754 | 787 | 415 | 427 | 419 | 427 | 419 |
| 1 | 4 | 1 | MTHFR | ASLT3 | 9450 | 36510 | 120 | 1135 | 1168 | 513 | 700 | 753 | 397 | 435 | 468 | 435 | 468 |
| 1 | 5 | 1 | MnSOD | ASLT4 | 9650 | 36510 | 120 | 978 | 981 | 456 | 693 | 740 | 419 | 285 | 288 | 285 | 288 |
| 1 | 6 | 1 | MnSOD | ASLT4 | 9850 | 36510 | 120 | 919 | 937 | 451 | 713 | 744 | 413 | 206 | 224 | 206 | 224 |
| 1 | 7 | 1 | MnSOD | ASLT5 | 10050 | 36510 | 120 | 1529 | 1633 | 772 | 695 | 733 | 390 | 834 | 938 | 834 | 938 |
| 1 | 8 | 1 | MnSOD | ASLT5 | 10250 | 36510 | 120 | 1298 | 1369 | 661 | 709 | 762 | 412 | 589 | 660 | 589 | 660 |
| 1 | 9 | 1 | VDR | ASLT39 | 10450 | 36510 | 120 | 12132 | 11284 | 3268 | 709 | 767 | 401 | 11423 | 10575 | 11423 | 10575 |
| 1 | 10 | 1 | VDR | ASLT39 | 10650 | 36510 | 120 | 10413 | 9749 | 3671 | 695 | 742 | 408 | 9718 | 9054 | 9718 | 9054 |
| 1 | 11 | 1 | VDR | ASLT40 | 10850 | 36510 | 120 | 22085 | 20350 | 7439 | 679 | 741 | 411 | 21406 | 19671 | 21406 | 19671 |
| 1 | 12 | 1 | VDR | ASLT40 | 11060 | 36510 | 120 | 21776 | 19840 | 7432 | 726 | 770 | 410 | 21050 | 19114 | 21050 | 19114 |

**Table 2**: A section of raw data table created by Genepix Pro 4.0 analysis software. X and Y show the specific location of the oligo spot on the slide. F635: gross feature (oligo spot) intensity at 635 nm wavelength. B635: background intensity at 635 nm wavelength.




| | | RD3619 |
|---|---|---|
| MTHFR | ASLT2 | 23953 |
| MTHFR | ASLT2 | 32911 |
| MTHFR | ASLT3 | 388 |
| MTHFR | ASLT3 | 360 |
| MnSOD | ASLT4 | 373 |
| MnSOD | ASLT4 | 490 |
| MnSOD | ASLT5 | 1275 |
| MnSOD | ASLT5 | 1806 |
| VDR | ASLT39 | |
| VDR | ASLT39 | |
| VDR | ASLT40 | 687 |
| VDR | ASLT40 | 774 |
| cyclinD1 | ASLT1 | 6657 |
| cyclinD1 | ASLT1 | 9061 |
| cyclinD1 | ASLT6 | 454 |
| cyclinD1 | ASLT6 | 564 |
| FABP2 | ASLT49 | |
| FABP2 | ASLT49 | |
| FABP2 | ASLT50 | |
| FABP2 | ASLT50 | |

| Gene | Genotype |
|---|---|
| MTHFR | Nt677 C/C |
| SOD | Nt-9 T/T |
| VDR | Cdn1 T/T |
| CyclinD1 | Cdn241 A/A |
| FABP2 | Cdn54 G/A |

**Figure 4**: Output of a SNParray hybridization. The signal intensities (left table) were obtained after processing of the raw data taking into consideration the background intensities. Genotypes (right table) were obtained on the basis of signal ratios.

**Development of SNParray Software (Task 4)**

We have designed a SNParray software to organize and evaluate the large amount of raw data (Table 2) obtained from SNParray scans. Manual analysis and evaluation of the raw data would have been very inefficient considering that approximately 1000 scans are proposed to be performed throughout the project. This SNParray software was designed to convert the net signal read by the analysis software, to different genotypes for each SNP. An algorithm was prepared first, and the software was written by the institute's microarray facility in PERL (a C based computer language).

The efficiency of each oligonucleotide, and therefore their signal intensity varies, providing a range of fluorescent signal intensities. A letter coding is assigned to different level of intensity values to distinguish the yield and the quality of the signal. Signals under certain values are considered as background. Software is able to read the signal-background values from the raw data table, and label each signal. Different genotypes for each SNP are defined for every possible letter combination in the algorithm, considering the defined cut-off values of all oligonucleotides. Each of the three genotype categories is subdivided into three perfection levels (excellent, good and poor) considering both the net signal intensity and the relative intensities of two allele specific oligonucleotide of the same SNP. The results that fall into the poor category will be repeated using a complementary TaqMan method for validation purposes (see below).

**Method Validation (Task 4)**

The SNParray method was validated for three selected SNPs using two additional methods that are routinely used in our laboratory. MTHFR 677C>T and MnSOD Val-9Ala polymorphims were validated using single strand conformation analysis (SSCP), and ER Ser10Ser SNP is validated using TaqMan, 5' nuclease allelic discrimination assay. The discrepancy rate for MnSOD SNP was a high with a frequency of about 10%. This was expected since the hybridization signal intensity of this specific SNP was very weak on the SNParray. The results of MTHFR and ER SNPs were highly concordant, with a discrepancy rate of less than 1%. The validation results for these three SNPs are provided in Figure 5.

**(a) ER Ser10Ser polymorphism**

| | | **TaqMan Results** | | | |
|---|---|---|---|---|---|
| | | CC | CT | TT | total |
| **SNParray results** | CC | 28 | | | 28 |
| | CT | | 70 | 1 | 71 |
| | TT | | | 26 | 26 |
| | total | 28 | 70 | 27 | 125 |

**(b) VDR Met1Thr polymorphism**

| SNParray results | SSCP Results: CC | CT | TT | total |
|---|---|---|---|---|
| CC | 46 | 1 | | 47 |
| CT | | 63 | | 63 |
| TT | | | 10 | 10 |
| total | 46 | 64 | 10 | 120 |

(c) **MnSOD Val-9Ala polymorphism**

| SNParray results | SSCP Results: CC | CT | TT | total |
|---|---|---|---|---|
| CC | 25 | 1 | | 26 |
| CT | 4 | 19 | 2 | 25 |
| TT | | | 20 | 20 |
| total | 29 | 20 | 22 | 71 |

**Figure 5**: Validation of genotypes of three SNPs (ER Ser10Ser, VDR Met1Thr and MnSOD Val-9Ala) using SNParrays, SSSC and TaqMan methods.

**Complementary Screening using TaqMan Method (Task 3 and Task 4)**

In order to genotype cases and controls for the SNPs that could not be studied using SNParrays we have developed a complementary TaqMan method. This method uses the Perkin-Elmer (PE) Applied Biosystems Sequence Detection 7900 HT System. This PCR based detection method uses allele-specific fluorescent probes, with a different label for each allele, to discriminate between alleles. Probes anneal in a sequence-specific (i.e. allele-specific) manner between the PCR primers, and in the course of the PCR the 5'-nuclease activity of the *Taq* polymerase releases the reporter dye of bound probes only, emitting an allele-specific fluorescence. The reporter fluorescent signal of probes is subdued by a quencher molecule in the intact probe, and does not release a signal. This methodology has the advantages of avoiding the use of restriction digests, hybridizations or electrophoresis thereby avoiding many sources of error and allowing high-throughput genotyping.

Using this approach we will have been optimizing the primers and probes using the Applied Biosystems-approved *Primer Express*™ probe and primer design program. The PCR conditions are applied according to the manufacturers protocol. Genotype analysis and calling is performed on amplified samples using the 7900HT software, using the standard procedures for automated allelic discrimination. In brief, by comparison to the fluorescence signals in known controls (homozygote allele 1, homozygote allele 2, and no

template), the software will call each "unknown" sample as homozygote allele 1, heterozygote, homozygote allele 2, undetermined, or no amplification (Figure 6). Outliers will be excluded as undetermined, and genotyping of these samples will be repeated.

We have currently optimized the conditions for six and completed the screening of three SNPs using this approach.

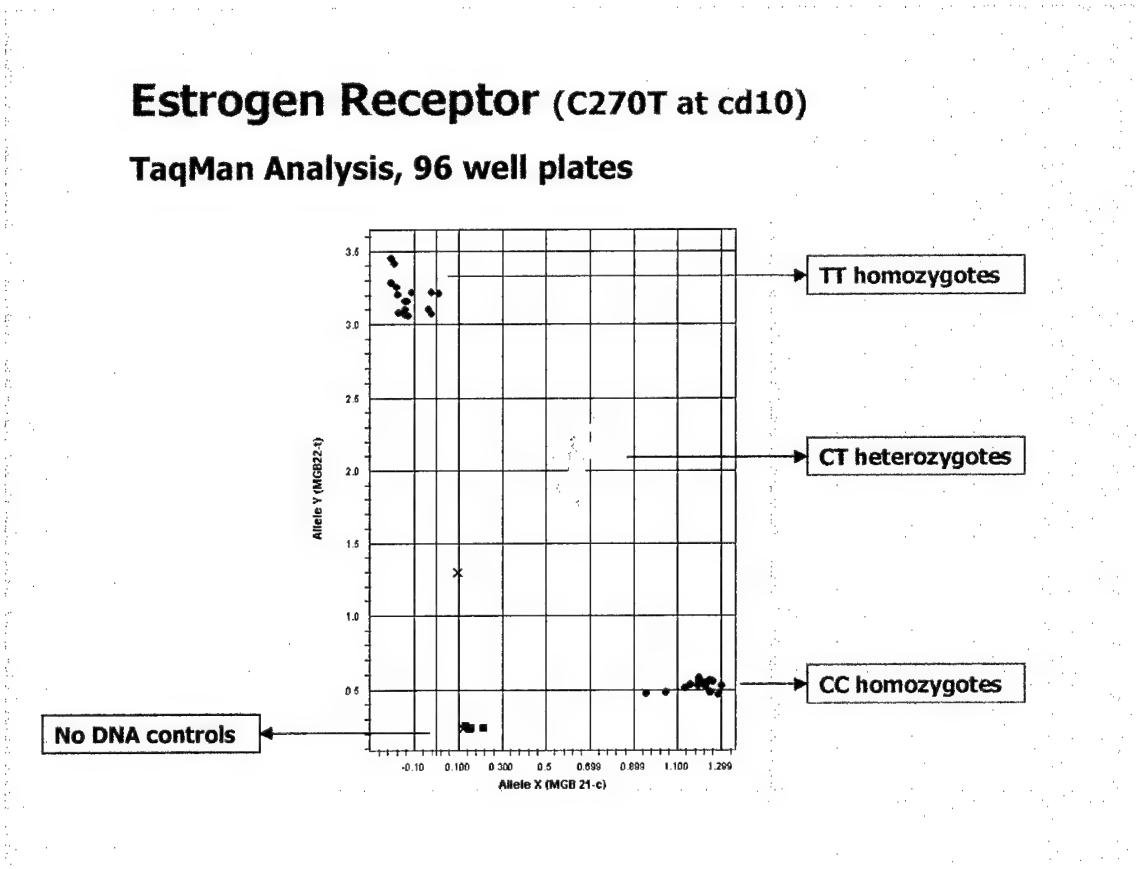



**Figure 6**: Output from Taqman assay. Genotypes (CC, TT and CT) are accumulated at different coordinates due to their signal intensity. This is done automatically by TaqMan software.

**Plating of Subject DNA Samples (Task 3 and Task 4)**

DNA samples of breast cancer cases and population controls are organized into 10 lists, each representing a 96-well micro-plate format. DNA samples are diluted to 5 ng/µl concentration, and transferred into 96-well master plates. Using an electronic 8-channel micropipette dispenser, numerous replica plates are prepared each well containing a total of 5ng of genomic DNA. Each replica plate is labeled, covered individually after all the DNA is air-dried, and stored at room temperature until they are used. An example for a plate organization is given in Figure 7. Each plate is prepared to contain three types of controls:

(a) Negative template controls; four negative DNA control samples are prepared in the same locations on each plate, controlling any possible DNA contamination.

(b) Experimental controls; 12 cell-line DNA samples (previously sequenced for all the listed SNPs) are located in the same coordinates in each plate. The results from these cell lines are used to evaluate the quality of probe preparation and hybridization.

(c) Validation controls; 10 DNA samples from each list are repeated in the following plate to evaluate the reproducibility of the results and to prevent DNA sample mix-up during master plate and/or replica plate preparations.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | Case1 | Case9 | Case16 | Case24 | Case32 | Case39 | Case47 | Case55 | Case62 | Case70 | Case77 | |
| B | Case2 | Case10 | Case17 | Case25 | Case33 | Case40 | Case48 | Case56 | Case63 | Case71 | Case78 | |
| C | Case3 | ND | Case18 | Case26 | Case34 | Case41 | Case49 | Case57 | Case64 | Case72 | Case79 | |
| D | Case4 | Case11 | Case19 | Case27 | Case35 | Case42 | Case50 | Case58 | Case65 | Case73 | Case80 | |
| E | Case5 | Case12 | Case20 | Case28 | ND | Case43 | Case51 | Case59 | Case66 | Case74 | | |
| F | Case6 | Case13 | Case21 | Case29 | Case36 | Case44 | Case52 | ND | Case67 | Case75 | | |
| G | Case7 | Case14 | Case22 | Case30 | Case37 | Case45 | Case53 | Case60 | Case68 | Case76 | | |
| H | Case8 | Case15 | Case23 | Case31 | Case38 | Case46 | Case54 | Case61 | Case69 | ND | | |

**Figure 7**: Design of subject DNAs and controls on the 96-well microplate.

**Screening of Breast Cancer Cases and Population Controls (Task 4)**

Currently we are screening cases and controls for 19 SNPs using the SNParray strategy described above. We are about to complete the results of the four out of total 10 plates. We are also currently optimizing the TaqMan assay conditions for the remaining SNPs which were unsuccessful with SNP arrays. Currently we have completed the screening of three out of 20 SNPs using TaqMan method.

**Immediate Future Task**

Using the SNP arrays and TaqMan method we will complete the screening of all the SNPs. Quality control and validation studies will be carried out throughout the screening to assess the reproducibility and accuracy. The results will be statistically analyzed to determine the association of SNPs with breast cancer risk.

**Key Research Accomplishments**

We have accomplished the tasks proposed in the Statement of Work by

- designing SNParray format and moving forward with their mass production
- optimizing the specificity of each probe individually on SNP arrays
- developing conditions for multiplex hybridization
- developing complementary genotyping method (TaqMan) for screening SNPs that were unspecific in with the SNParray protocol
- developing a complicated software to organize and evaluate the raw data obtained from SNParray scans
- Validating the SNParray method with two other genotyping methods
- preparing and plating all the subject DNA samples to be used in the study
- screening a portion of the study cases and controls

**Reportable Outcomes**

This work has been presented at following conferences.

1. Venus Onay, Julia Knight, and Hilmi Ozcelik, "Candidate SNP Analysis in the Study of Breast Cancer Risk Using SNParrays." Controversies in the Etiology, Detection and Treatment of Breast Cancer:2002, June 13-14,2002, Toronto, Ontario, Canada.

2. Venus Onay, Julia Knight, and Hilmi Ozcelik, "Candidate SNP Analysis in the Study of Breast Cancer Risk Using SNParrays." 93rd Annual Meeting of AACR, April 6-10,2002, San Francisco, California, USA.

3. Venus Onay, Julia Knight, and Hilmi Ozcelik, "Identifying the Role of SNPs in Breast Cancer Risk Using Microarray Technology." Oncogenomics Conference, 25-27 January 2001, Tucson, Arizona, USA.

4. Venus Onay, Julia Knight, and Hilmi Ozcelik, "Microarray Technology to Study the Role of Candidate SNPs in Breast Cancer Risk" 3rd International Meeting on Single Nucleotide Polymorphism and Complex Genome Analysis, 8th-11th September 2000, Taos, New Mexico, USA.

## References


1. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavigain S, Liu Q, Cochran C, Bennet LM, Ding W, Bell R, Rosenthal J, Hussey C, Tran T, McClure M, Frye C, Hattier T, Phelps R, Haugen-Strano A, Katcher H, Yakumo K, Gholami Z, Shaffer D, Stone S, Bayer S, Wray C, Bogden R, Dayananth P, Ward J, Tonin P, Narod S, Bristow PK, Norris FH, Helvering L, Morrison P, Rosteck P, Lai M, Barret C, Lewis C, Neuhausen S, Cannon-Albright L, Goldgar D, Wiseman R, Kamb A, and Skolnick M.(1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 266: 66.

2. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G.(1995) Identification of the breast cancer susceptibility gene BRCA2. Nature 21:789.

3. Tavtigian SV, Simard J, Rommens J, Couch F, Shattuck-Eidens D, Neuhausen S, Merajver S, Thorlacius S, Offit K, Stoppa-Lyonnet D, Belanger C, Bell R, Berry S, Bogden R, Chen Q, Davis T, Dumont M, Frye C, Hattier T, Jammulapati S, Janecki T, Jiang P, Kehrer R, Leblanc JF, Goldgar DE, et al. (1996) The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. Nature Genetics 12(3):333.

4. Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen N, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lander ES, et al (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280(5366):1077

5. Collins F, Patrinos A, Jordan E, Chakravarti A, Gesteland R, Walters L, the members of the DOE and NIH planning groups New Goals for the U.S. Human Genome Project:1998-2003. Science 282: 682

6. Hacia JG, Brody LC, Chee MS, Fodor SP, Collins FS (1996) Detection of heterozygous mutations in BRCA1 using high density oligonucleotide arrays and two-colour fluorescence analysis. Nat Genet 14(4):441.

7. Fernandez PL, Jares P, Rey MJ, Campo E, Cardesa A (1998) Cell cycle regulators and their abnormalities in breast cancer. Mol Pathol 51(6):305

8. Blais A, Labrie Y, Pouliot F, Lachance Y, Labrie C (1998) Structure of the gene encoding the human cyclin-dependent kinase inhibitor p18 and mutational analysis in breast cancer. Biochemistry and Biophysical Research Communications 247, 146.

9. Zariwala M, Xiong Y (1996) Lack of mutation in the cyclin-dependent kinase inhibitor, $p19^{INK4d}$, in tumor-derived cell lines and primary tumors. Oncogene 13, 2033.

10. Ferrando A, Balbin M, Pendás AM, Vizoso F, Velasco G, López-Otin C (1996) Mutational analysis of the human cyclin-dependent kinase inhibitor $p27^{kip1}$ in primary breast carcinomas. Hum Genet 97, 91.

11. Li, Y, Millikan RC, Newman B, Conway K, Tse C-KJ, Liu ET. (1999) P57 (KIP2) polymorphisms and breast cancer risk. Hum Genet 104, 83.

12. Mousses S, Özçelik H, Lee PD, Malkin D, Bull SB, Andrulis IL (1995). Two variants of the CIP1/WAF1 gene occur together and are associated with human cancer. Human Molecular Genetics 4, 1089.

13. Hosokawa, Y, Tu T, Tahara H, Smith AP, Arnold A (1995) Absence of cyclin D1/PRAD1 point mutations in human breast cancers and parathyroid adenomas and identification of a new cyclin D1 gene polymorphism. Cancer Letters 93, 165.

14. Matthias C, Branigan K, Jahnke V, Leder K, Haas J, Heighway J, Jones PW, et al. (1998) Polymorphism within the Cyclin D1 gene is associated with prognosis in patients with Squamous Cell Carcinoma of the head and neck. Clinical Cancer Research 4, 2411.

15. Wang-Gohrke S, Rebbeck TR, Besenfelder W, Kreienberg R, Runnebaum IB (1998). P53 germline polymorphisms are associated with an increased risk for breast cancer in German women. Anticancer Research 18, 2095.

16. Mavridou D, Gornall R, Campbell IG, Eccles DM (1998) TP53 intron 6 polymorphism and the risk of ovarian and breast cancer. British Journal of Cancer 77, 676.

17. Blaszyk H, Hartmann A, Sommer SS, Kovach JS. (1996) A polymorphism but no mutations in the GADD45 gene in breast cancers. Hum Genet 97, 543.

26. Carroll BT, Couch FJ, Rebbeck TR, Weber BL. (1999) Polymorphisms in PTEN in breast cancer families. J Med Genet 36, 94.

18. Thai TH, Du F, Tsan JT, Jin Y, Phung A, Spillman MA, Massa HF, et al. (1998). Mutations in the BRCA1-associated RING domain (BARD1) gene in primary breast, ovarian and uterine cancers. Human Molecular Genetics 7, 195.

19. Rutter JL, Mitchell TI, Butticè G, Meyers J, Gusella JF, Ozelius LJ, Brinckerhoff CE. (1998) A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter creates an Ets binding site and augments transcription. Cancer Research 58, 5321.

20. Dybdahl M, Vogel U, Frentz G, Wallin H, Nexø BA. (1999) Polymorphisms in the DNA repair gene XPD: Correlations with risk and age at onset of Basal Cell Carcinoma. Cancer Epidemiology, Biomarks & Prevention 8, 77.

21. Lehrer SP, Schmutzler RK, Rabin JM, Schachter BS. (1993) An estrogen receptor genetic polymorphism and a history of spontaneous abortion –Correlation in women and estrogen receptor positive breast cancer but not in women with estrogen receptor negative breast cancer or in women without cancer. Breast Cancer Research and Treatment 26, 175.

22. Iwase H, Greenman JM, Barnes DM, Hodgson S, Bobrow L, Mathew CG. (1996) Sequence variants of the estrogen receptor (ER) gene found in breast cancer patients with ER negative and progesterone receptor positive tumors. Cancer Letters 108, 179.

23. Hirvonen A (1999) Polymorphisms of Xenobiotic-Metabolizing Enzymes and Susceptibility to Cancer. Environ Health Perspect 107 Suppl 1:37.

24. Lai C & Shields PG (1999) The role of interindividual variation in human carcinogenesis. J Nutr 129(2S Suppl):552S.

25. Nebert DW, McKinnon RA, Puga A. (1996) Human drug-metabolizing enzyme polymorphisms:Effects on risk of toxicity and cancer. DNA and Cell Biology 15, 273.

26. Feigelson HS, Coetzee GA, Kolonel LN, Ross RK, Henderson BE. (1997) A polymorphism in the CYP17 gene increases the risk of breast cancer. Cancer Research 57, 1063.

27. Helzlsouer KJ, Selmin O, Huang H-Y, Strickland PT, Hoffman S, Alberg AJ, Watson M, et al. (1998). Association between glutathione S-transferase M1, P1 and T1 genetic polymorphisms and development of breast cancer. Journal of the National Cancer Institute 90, 512.

28. Inskip A, Elexperu-Camiruaga J, Buxton N, Dias PS, MacIntosh J, Campbell D, Jones PW, Yengi L, Talbot JA, Strange RC, et al (1995) Identification of polymorphism at the glutathione S-transferase, GSTM3 locus: evidence for linkage with GSTM1*A. Biochem J 312:713.

29. Thompson PA, Shields PG, Freudenheim JL, Stone A, Vena JE, Marshall JR, Graham S, et al. (1998) Genetic polymorphisms in catechol-*O*-methyltransferase, menopausal status, and breast cancer risk. Cancer Research 58, 2107.

30. Robarts AH. A review of intestinal fatty acid binding protein gene variation and the plasma lipoprotein response to dietary components (1998). Clin Biochemistry 31(8);609.

31. Chen J, Giovannucci EL, Hunter DJ. (1999) MTHFR polymorphism, methyl-replete diets and the risk of colorectal carcinoma and adenoma among U.S. men and women: An example of gene-environment interactions in colorectal tumorigenesis. J Nutr 129, 560S.

32. Ambrosone CB, Freudenheim JL, Thompson PA, Bowman E, Vena JE, Marshall JR, Graham S, et al. (1999) Manganese superoxide dismutase (MnSOD) genetic polymorphisms, dietary antioxidants, and risk of breast cancer. Cancer Research 59, 602.

33. Tartour E & Fridman WH (1998) Cytokines and cancer. Int Rev Immunol 16(5-6):683

34. Maini A, Morse PD, Wang CY, Jones RF, Haas GP (1997) New developments in the use of cytokines for cancer therapy. Anticancer Res 17(5B):3803.

35. Salgaller ML & Lodge PA (1998) Use of cellular and cytokine adjuvants in the immunotherapy of cancer. J Surg Oncol68(2):122.

36. Crichton MB, Nichols JE, Zhao Y, Bulun SE, Simpson ER (1996) Expression of transcripts of interleukin-6 and related cytokines by human breast tumors, breast cancer cells, and adipose stromal cells. Mol Cell Endocrinol 118(1-2):215.

37. Douglas AM, Goss GA, Sutherland RL, Hilton DJ, Berndt MC, Nicola NA, Begley CG (1997) Expression and function of members of the cytokine receptor superfamily on breast cancer cells. Oncogene 1997 14(6):661.

38. Kurtzman SH, Anderson KH, Wang Y, Miller LJ, Renna M, Stankus M, Lindquist RR, Barrows G, Kreutzer DL (1999) Cytokines in human breast cancer: IL-1alpha and IL-1beta expression. Oncol Rep 6(1):65.

39. Kroeger KM, Carville KS, Abraham LJ. (1997) The –308 tumor necrosis factor-α promoter polymorphism effects transcription. Molecular Immunology 34, 391.

40. Hurme M, Lahdenpohja N, Santtila S. (1998) Gene polymorphisms of interleukins 1 and 10 in infectious and autoimmune diseases. Ann Med 30,469.

41. Knight J.A, Marrett L.D., Weir H. Occupation and Risk of Germ Cell Testicular Cancer by Histologic Type in Ontario. Journal of Occupational and Environmental Medicine. 1996; 38(9):884.